A TOPOLOGICAL VIEW OF SOME PROBLEMS IN COMPLEXITY THEORY

Michael Sipser
Mathematics Department
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

ABSTRACT.

We present a new, combinatorial proof of the classical theorem that the analytic sets are not closed under complement. Possible connections with questions in complexity theory are discussed.

INTRODUCTION.

A number of recent results in circuit complexity theory have been stimulated by a new understanding of certain old theorems in descriptive set theory. The hierarchy theorem for polynomial-size, constant depth circuits is the finite counterpart to the Borel rank hierarchy theorem [S2]. The lower bound for circuits computing the parity function [FSS, A] in part stemmed from a result showing that infinite parity functions are not Borel definable [S1]. In both cases, the classical proofs do not exhibit enough combinatorial structure to yield insight into the finitary questions and new proofs were required.

The link between circuits and Borel sets stems from an analogy between polynomial growth and countability [S1]. In this paper, we propose a further link suggested by this analogy, one between NP and the analytic sets [K, M]. Several observations support this connection. NP sets are exactly those which are accepted by polynomial-size, nondeterministic circuits (ignoring uniformity issues). A Nondeterministic circuit is one with inputs that are nondeterministically set as well as ordinary inputs. By the addition of additional nondeterministic inputs these circuits may be converted to equivalent polynomial-size, depth-2, nondeterministic circuits. The infinitary analog to these, the countable, depth-2, nondeterministic circuits accept exactly the analytic sets.

This analogy suggests that the NP = co-NP question may be illuminated by the theorem stating that the class of analytic sets is not closed under complementation. The classical proof by diagonalization of this theorem does not seem to have a corresponding finitary argument. We give here a new purely combinatorial proof of this theorem.

PRELIMINARIES.

Let $\Sigma = \{0,1\}$ and Σ^{ω} be the set of infinite 0,1 sequences or reals. An interval is the set of reals extending a finite sequence. An open set is a union of intervals. Closing the open sets under countable union and intersection gives the Borel sets. An analytic set is a projection of a Borel set, (i.e., A is analytic if $A = \{\alpha : \langle\alpha,\beta\rangle \in B$ for some $\beta\}$ where B is Borel and $\langle\alpha,\beta\rangle$ is any pairing function).

Definition. A literal is a member of $\{x_1, \bar{x}_1, x_2, \bar{x}_2, \ldots\}$. A $\bigvee_1$-circuit is a collection of literals and an $\bigwedge_2$-circuit is a countable collection of V_1-circuits. These naturally represent functions from Σ^ω to Σ. A <u>nondeterministic circuit</u> has additional nondeterministic inputs represented by literals drawn from $\{y_i,\overline{y_i}\}$. It accepts a given real if there is some setting of the nondeterministic inputs which causes evaluation to 1. If a nondeterministic Λ_2-circuits accepts a real ρ, then a setting π of the x and y inputs causing evaluation to 1 is called a <u>proof</u>. If C is a member V_1-circuit then π <u>satisfies C at j</u> if the j^{th} literal of C is 1 in π.

The nondeterministic circuits accept exactly the class of analytic sets.

Let $N = \{1,2,\ldots\}$ and N^* be the set of finite sequences over N. A <u>tree</u> is a subset of N^* closed under prefix. Let T be the set of all trees. We fix any enumeration of N^* and obtain a natural correspondence between trees and reals. Hence we may speak of, say, an analytic set of trees. A tree is <u>well-founded</u> if it has no infinite branch, (i.e., tree τ is well-founded every $\alpha \in N^\omega$ contains a prefix $b \notin \tau$). Let W be the set of all well-founded trees. It is easy to verify that $\overline{W}$, the complement of W, is analytic. (nondeterministically guess the branch). We show that W itself is not.

We introduce some additional notation. If s,t are sequences in N^* then st is the concatenation of s and t. If A is a set of sequences then $sA = \{st: t \in A\}$.

The Proof.

Theorem. There is an analytic set $\overline{W}$ whose complement is not analytic.

Proof. Let W be the set of all well-founded trees. We first establish the following Ramsey-like property of collections of trees.

Definition. For any tree τ, collection of trees A, and $s \in N^*$ we say the <u>detail of τ at s,</u> $\tau^s = \{t: st \in \tau\}$. The <u>detail of A at s</u>, $A^s = \{\tau^s: \tau \in A\}$. Say A <u>is large at s</u> if $W \subseteq A^s$ or simply <u>large</u> if it is large for some s. For example, W is large at e, the sequence of length o.

Claim. If A is large at s and is divided into a countable union of sets, $A = B_1 \cup B_2 \cup \ldots$ then for some i and j, B_i is large at sj.

Proof. Assume to the contrary that for each i,j B_i is not large at sj. So each detail of B_i at any sj lacks a tree $\tau_{i,j}$ in W. By pasting these together, one obtains the well-founded tree $\sigma = 1\tau_{1,1} \cup 2\tau_{2,2} \cup \cdots$ not in B_i^s for any i and therefore not in A^s. But $\sigma \in W$ contradicting the largeness of A at s.

□

To show that W is not analytic, we construct a sequence of large sets $W \supseteq A_1 \supseteq A_2 \supseteq \cdots$ containing trees which "converge" to one not in W.

Assume to the contrary that W is analytic, accepted by a nondeterministic Λ_2-circuited N containing V_1-circuits $C_1, C_2, \ldots$. Let $N^* = \{t_1, t_2, \ldots\}$. We perform a construction in stages. The goal of stage i is to construct $A_i \subseteq W$, $s_i \in N^*$, and $p_i \in N^*$ such that A_i is large at s_i, all $\sigma \in A_i$ agree on $t_1, \ldots, t_i$, and each $\sigma \in A_i$ has a proof which satisfies C_j at $p_i(j)$

(the j^{th} position of p_i) for $j \leq i$. Let $A_0 = W$, $s_i = e$, and $p_i = e$. Go to Stage 1.

<u>Stage i.</u> Let $B_m = \{\alpha \in A_{i-1}: \alpha$ has a proof which satisfies C_i at $m\}$. By the lemma, for some m and n, B_m is large at $s_{i-1}n$. Fix m and n. Let $D = \{\alpha \in B_m: \alpha$ contains $t_i\}$ and $E = B_m - D$. By the lemma, either D or E is large at a sequence $s_{i-1}nk$. Let A_i be this large set. Let $s_i = s_{i-1}nk$ and $p_i = p_{i-1}m$. Go to stage i+1.

It is straightforward to verify that upon completion of all stages there is exactly one tree α in every A_i. Furthermore α is not in W since it contains an infinite branch $S_1 \cup S_2 \cup \cdots$ and there is a proof $\pi = p_1 \cup p_2 \cup \cdots$ which satisfies every C_i. Therefore α is accepted by N, a contradiction.

<u>CONCLUSION.</u>

The links between topological notions such as open set, Borel set, and analytic set and their companions in circuit complexity bear further investigation. Ajtai's theorem [A] that every polynomial-size, depth-k definable set is well approximable by a union of cylinders is analogous to the theorem that all Borel sets are measurable, i.e., well approximable by open sets. There seems to be a parallel between Baire category theorem type constructions and constructions involving probabilistic methods. It is interesting to view these observations in the context of defining a notion of finite topological space.

ACKNOWLEDGEMENT.

I am deeply grateful to John Addison for having introduced me to descriptive set theory.

REFERENCES

[A] M. Ajtai, "Σ_1^1-formulae on finite structures", Annals of Pure and Applied Logic 24, 1983, 1-48.

[FSS] M. Furst, J.B. Saxe, M. Sipser, "Parity, circuits, and the polynomial time hierarchy", Proceedings of the 22nd Annual Symposium on Foundations of Computer Science, 1981, 260-270.

[K] K. Kuratowski, Topology, Academic Press, 1966.

[M] Y. Moschavakis, Descriptive Set Theory, North-Holland, 1980.

[S1] M. Sipser, "On polynomial versus exponential growth", unpublished report, 1981.

[S2] M. Sipser, "Borel sets and circuit complexity", Proceedings of the 15th Annual Symposium on Theory of Computing, 1983, 61-69.